The efficacy of hiring credits in distressed areas

Jorge Pérez* and Michael Suher†

April 2017
Preliminary and incomplete. Please do not cite.

Abstract We analyze the efficacy of hiring tax credits, particularly in distressed labor markets. These types of programs have proven hard to assess as their introduction at the state level tends to be endogenous to local conditions and future prospects. We conduct an empirical study of a hiring tax credit program implemented in North Carolina in the mid 1990s, which has a quasi-experimental design. Specifically, the 100 counties in the state are ranked each year by a formula trying to capture their economic distress level. The generosity of the tax credits jumps discontinuously at various ranking thresholds. We estimate the impact of the credits using difference in differences and regression discontinuity methods. Our estimates show positive impacts on employment levels of around 3% and fairly sizable impacts on unemployment - a $10,000 credit leads to a 0.7 percentage point reduction in the unemployment rate.

*Department of Economics, Box B. Brown University. Providence RI 02912 (e-mail: jorge_perez@brown.edu)

†Furman Center for Real Estate and Urban Policy, New York University, 139 MacDougal St., New York, NY 10012 (e-mail: michael.suher@nyu.edu)

1 Introduction

Hiring tax credits are a commonly used tool at the state level and as part of federal programs to address both short run downturns and longer run economic distress. They are place-based in that they aim to revitalize a specific geography rather than individual workers. Prior evidence on their efficacy has been mixed, with zero or small positive impacts on employment seen depending on the type of credit.

The empirical evaluation of these policies is difficult as their enactment is typically designed to be endogenous to expected economic prospects or local economic distress. The direction of the bias is also not clear. Significantly poor economic performance may swamp estimates of program impact even if they are positive. Alternatively, natural mean reversion in areas which recently experienced negative shocks could be incorrectly attributed to the policy intervention. We examine a series of tax credit programs enacted in the state of North Carolina whose structure enables causal estimates of policy impacts. The programs include thresholds at which credit size jumps discontinuously allowing for difference in differences (DD) as well as regression discontinuity (RD) estimates. Our DD estimates show zero or occasional negative impacts on employment. The RD estimates are noisy but show increases of around 3%, indicating the importance of accounting for time varying unobservables. We find significant reductions in unemployment rates which grow over time, and are on the order of 0.7 percentage points for a $10,000 credit.

A few papers have considered the theoretical case of subsidizing hiring. Neumark (2013) highlights the potential outsize benefit of policies like hiring credits aimed at stimulating labor demand during a recession, where downward wage rigidities will mute the impact of labor supply policies. Kline and Moretti (2013) augment a spatial equilibrium model with persistent long run differences in unemployment rates across areas, as is empirically observed. Firms in high unemployment/low productivity areas post too few vacancies due to excessive hiring costs providing a rationale for subsidizing hiring in distressed areas. Amior and Manning (2015) find that distressed areas experience serially correlated negative demand shocks that lead to swift and continual outflows of employment. Population leaves as well, but not as fast as employment, leading unemployment rates to remain elevated.

Bartik (2001) notes that when made generally available, a large share of credits ends up as wastage, meaning they are claimed by firms whose hiring would have taken place absent the program. Credits are typically small enough though that they can still compare favorably to other employment subsidy policies even after accounting for the wastage. The use of credits during recessions and/or in distressed areas can reduce the concerns of wastage, though churning may still be an issue. This is where firms increase hiring while simultaneously increasing separations. Because our analysis focuses on overall employment levels and unemployment rates, estimates of program effects will be net of any churning.

Neumark and Grijalva (2015) use cross-state variation in the adoption of hiring credits to estimate their impact. They find no impacts on employment growth in general but small positive effects during recessionary periods or when programs incorporate recapture provisions. They do not consider the size of the credits though. In order to deal with the endogeneity of the adoption of these programs, they rely on projected counterfactual employment trends based on state industrial composition prior to enactment of a policy. While such predicted employment measures are known to correlate strongly with actual employment changes at decadal frequency, they may fail to capture shorter run counter-factual employment trends, particularly during recessions. Our use of a regression discontinuity design can arguably offer more causal estimates of program effects. We also extend the analysis by considering the impact of hiring credits on unemployment rates.

Chirinko and Wilson (2016) also exploit cross-state variation in hiring credit adoption with a focus on the potential for fiscal foresight, wherein programs that are pre-announced could lead firms to initially depress hiring and then ramp up once the credits become available. They find evidence of positive impacts on employment at a lag of two to three years, consistent with our findings. They also find pre-program dips, which can upwardly bias estimates of program effects by 33%. Our reliance on annual rather than monthly employment levels should help alleviate this bias.

Busso et al. (2013) and Freedman (2013) both assess the impact of the federal enterprise zone program which includes hiring credits, and find positive employment effects for neighborhood residents. The policy we investigate differs in that it targets a labor market rather than a specific neighborhood.

Our study builds on this prior research on the efficacy of hiring tax credits by incorporating plausibly causal regression discontinuity estimates at the labor market level and considering the size of credits made available. We also look at unemployment rates in addition to the conventional focus on employment growth.

We describe the mechanics of North Carolina's hiring tax credits in section 2. In section 3, we summarize our data sources and give an overview of the labor market during our sample period. We describe the estimation strategy in section 4. Section 5 presents the estimation results. Section 6 concludes.

2 North Carolina's hiring tax credit programs

In the mid 1980s, North Carolina government officials were concerned with the divergence in economic fortunes among the state's 100 counties. A tax incentive program began in 1988 to address the situation in the least economically robust counties. The state Department of Commerce was tasked with ranking counties each year from 1 to 100 based on economic distress, which legislation defined as the combination of a high unemployment rate and low per capita incomes. Businesses in the 20 most distressed counties were then eligible for a $2,800 tax credit for each new full time employee hired. The number of eligible counties was progressively increased, and had reached 50 by the time the program ended in 1995. In its place, a revamped program was launched in 1996 known as Article 3A or the William S. Lee program. It continued the use of a county ranking scheme, but extended tax credits to all counties. Further, it grouped counties into five tiers based on the distress rankings with larger credits of $12,500 available to the 10 lower tier/more distressed counties. Firms in less distressed counties could receive credits between $500 and $4,000 per new hire. Over the course of this program, the number of counties eligible for the largest credit size was increased as low population and high poverty rates were added as overrides of the distress ranking system, with 28 counties eligible for the largest credits by the final year of the program in 2006. The William S. Lee program was itself replaced in 2007 by the Article 3J program. This latter program operated in similar fashion to is predecessor, but with some changes to the credit eligibility formulas. Now there would be only three tiers, and tier size would remain fixed. Tier 1 - the most distressed - contained 40 counties eligible for credits of $12,500. The 40 counties in tier 2 could receive $5,000 credits and the highest performing 20 counties in tier 3 could receive $750 credits. The distress ranking formula was also amended to incorporate population growth and property value per capita alongside unemployment rates and income. In 2014, the Article 3J program ended and was not replaced (Program Evaluation Division, 2015).

In this study we focus on the William S. Lee program which began in 1996, and is denoted as wave 1 below, as it provides the cleanest quasi-experimental set-up. Because counties were re-ranked every year, treatment status was not always constant, with occasional slippage between tiers, in addition to the legislated expansion of the lowest tier over time. While there was no requirement that hires be of certain types of workers, such as those currently unemployed, it was restricted based on industry. The main industries eligible were manufacturing, wholesale trade, warehousing, and those related to data processing.

Building 1 to 100 rankings using a somewhat ad-hoc choice of inputs meant high performing counties often received lower tier status than clearly more distressed counties. This became even more pronounced once small population overrides to the rankings were introduced. A more continuous and robust measure of distress was proposed by the state Department of Commerce though not adopted (Department of Commerce, 2014). Below we demonstrate that lower ranked counties do have lower population growth, higher poverty, and lower per capita income. There is no evidence of discontinuities in pre-treatment conditions at the program thresholds though. The ranking variable is also not strongly correlated with post-treatment outcomes after controlling for tier status, allowing us to do comparisons between counties farther away from the cut-offs.

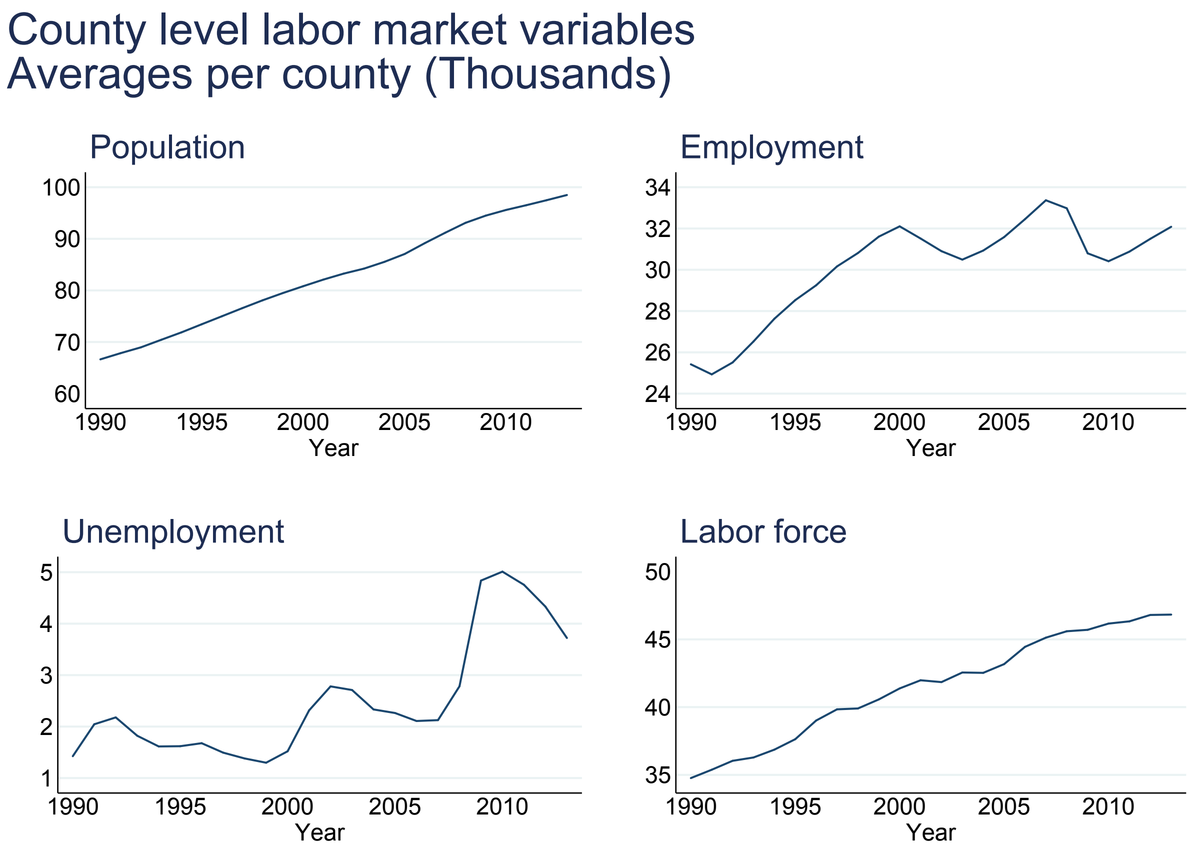
3 Data

Employment, unemployment, and labor force data is from the Bureau of Labor Statistics (BLS). Hiring and separations data come from the Census Bureau's Quarterly Workforce Indicators (QWI). Tier status comes from annual reports issued by the state of North Carolina and archived versions of the Commerce Department's website. The rankings are reconstructed using data from the BLS, population and income data from the Bureau of Economic Analysis, and poverty data from the Census and the United States Department of Agriculture, in conjunction with the rules laid out in the legislation creating and amending the programs. Our sample period runs from 1990 to 2013. Figure 1 shows overall conditions in North Carolina's economy during the program period. Population is steadily growing throughout. Employment is growing through 2000. It fluctuates through the 2000s, and returns to about the level of 2000 by the end of the period.

Figure 2 looks at the evolution of employment and unemployment at the county level, dividing the state into three groupings which roughly approximate the three tier groups for the program analysis[1]. The distress rank is correlated with these economic indicators. The counties ranked least distressed (highest ranks) have much higher employment to population ratios, while the middle and lower rank groups are very similar on this measure. All three groups are more clearly separated by unemployment rate which is by design. Most importantly though for our analysis, the evolution of employment and unemployment rates is largely parallel across rank groups.

[1]We use a constant group size here to portray trends separately from the compositional change due to the legislated expansion of the tier 1 group over time.

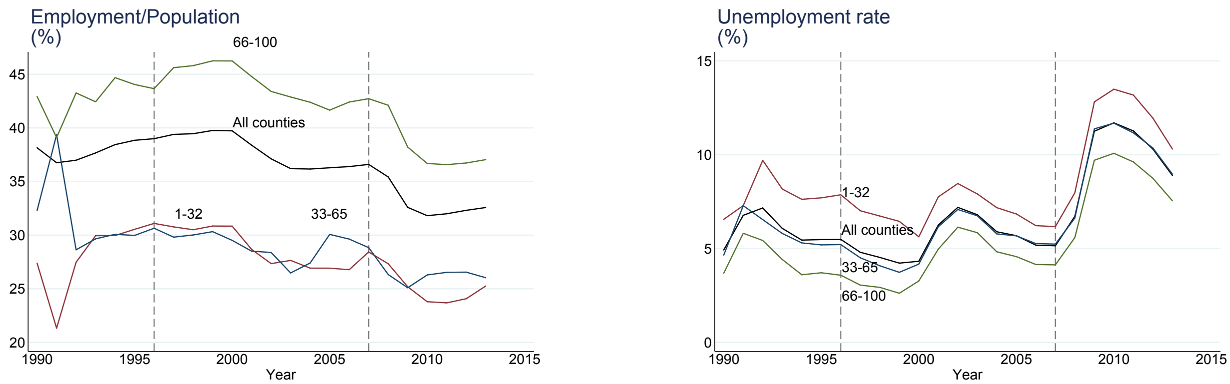
Figure 1



Note: Population data is from the Bureau of Economic Analysis. Employment, unemployment, and labor force data is from the Bureau of Labor Statistics.

Figure 2: Evolution of employment and unemployment outcomes across counties with different distress ranks.

(a) (b)



Note: Lines depict averages across counties classified according to their economic distress rank. The composition of the groups changes over time as counties change rank. Vertical lines denote the beginning of the first and second waves of the program.

4 Estimating the effect of hiring credits

In this section we describe our estimation strategy. We lay out the difficulties of estimating the effect of hiring credits and suggest two strategies that take advantage of the assignment of subsidies based on distress ranks. First, we outline a difference in differences estimation strategy that compares counties across different tiers. Then, we describe a regression discontinuity design strategy that exploits changes in subsidy status across distress ranking thresholds.

4.1 Difference in differences estimates

We group North Carolina's 100 counties into three groups defined by the subsidy program tiers. For the first wave of the program, running from 1996 to 2007, group 1 contains the 10 to 28 most distressed counties that receive the highest subsidy amount, $12,500. Group 2 contains the next 24 to 42 counties that receive either 4 or 3 thousand dollars, and group 3 contains the remaining 30 counties. We label these groups tier 1, tier 2 and tier 3 in our analysis.[2]

Figure 3 shows the relationship between several variables and the economic distress rank.

[2]Under the official program definitions, there are five tiers in the first wave of the program. Tier 1 is the same as our definition. We combine tiers 2 and 3 together and tiers 4 and 5 together as they have similar program intensity as measured by the credit size for which they are eligible.

Two messages emerge from these figures. As expected, there is an overall negative relationship between economic outcomes and economic distress. Poverty rates are higher for more distressed counties, while income per capita and population growth are lower. However, this relationship is smooth across the tier cutoffs. Moreover, the slope of this relationship is small, suggesting that the economic distress rank is not strongly correlated with these outcomes, within each tier and across tiers for counties with similar ranks.

We estimate the effect of the program by comparing the evolution of employment and unemployment across counties in different tiers. To avoid making comparisons between extremely different counties, we only do comparisons across consecutive tiers. Tier 3 contains counties with major cities, which may have very different dynamics compared to small distressed counties. To estimate the effect of receiving the largest subsidy, we focus on comparing counties in tier 1 to those in tier 2. Our basic specification is

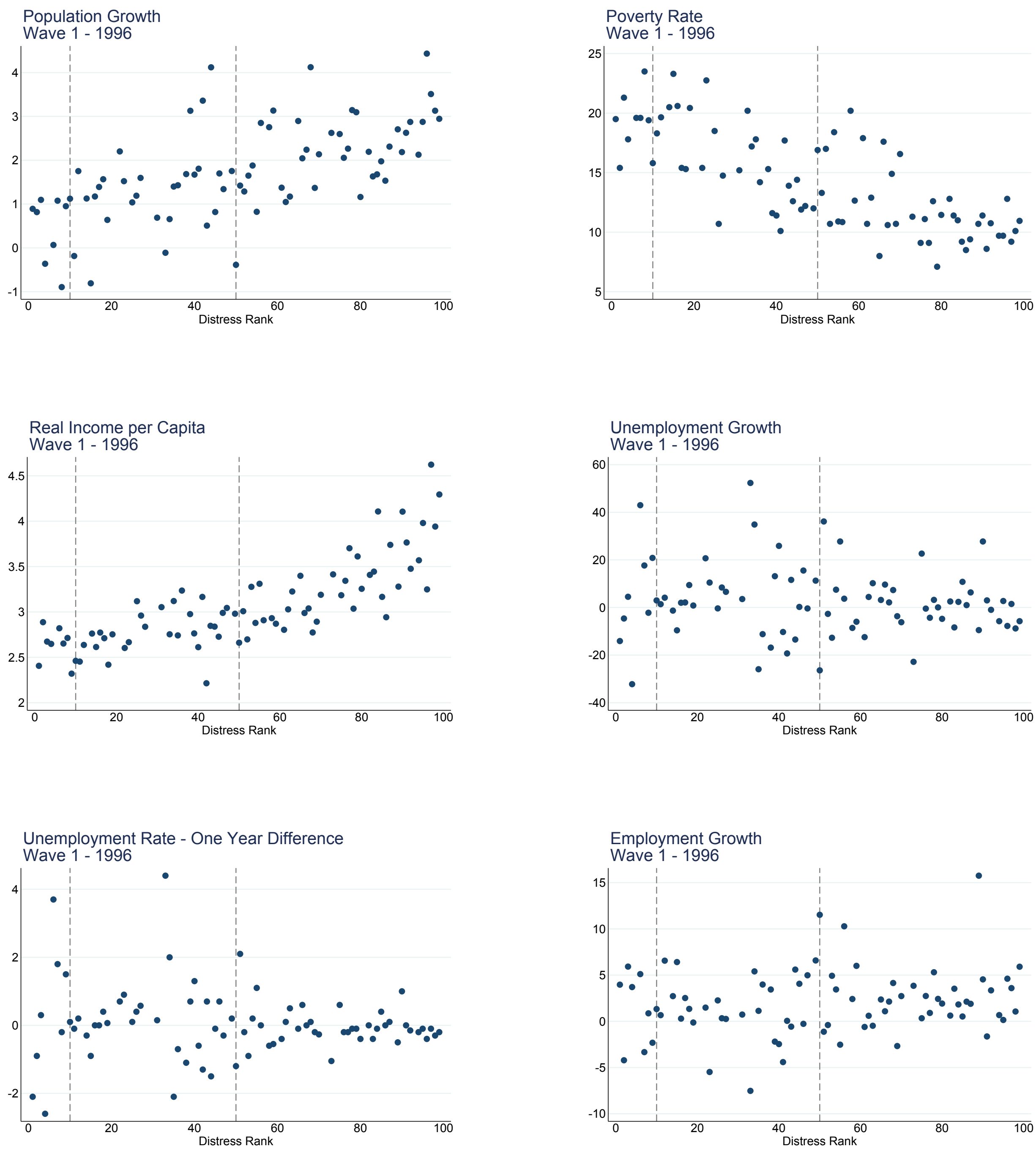
$$Y_{ct} = \beta_0 + \gamma_c + \gamma_t + \sum_{k=0}^{K} \theta_k tier1_{t-k} + \beta X_{ct} + \varepsilon_{ct} \tag{1}$$

Here, Y_{ct} is the outcome of interest for county c at time t. γ_c and γ_t are county and year effects intended to capture permanent differences across counties and common shocks that affect all the counties in each year. $tier1_{t-k}$ is a dummy variable equal to 1 whenever a county is assigned to tier 1. The coefficients of interest, θ_k capture the contemporaneous and lagged effects of tier 1 status on the outcome variables. Allowing for lagged effects is essential since the hiring subsidy programs may take a few years to gain traction and have a noticeable effect on employment (Neumark and Grijalva, 2015). Moreover, including lagged effects allows us to compare counties that have a similar history of treatment. By looking at the differences in these coefficients, we can assess how the effect of the program changes over time. We estimate this specification using only counties in tiers 1 and 2, from 1996 to 2007.

This difference in differences strategy is valid only if unobservables have a similar evolution over time across tiers. We validate whether this is the case by looking at the evolution of the outcomes across different tiers. We also include control variables X_{ct} to allow for some county heterogeneity. We include lags of order $K + 1$ of population and real income per capita, as well as lags of employment and labor force variables. This addresses the possibility of counties evolving heterogeneously because of different initial conditions before the beginning of the program. It also addresses mean reversion in outcomes (Heckman et al., 1999). We also allow for tier and county-specific linear time-trends.

We calculate clustered standard errors by county to allow for serial correlation in the error term ε_{ct} within counties (Bertrand et al., 2004). Some of our regressions have less than

Figure 3: Pre-program outcomes



Note: County level economic indicators are arrayed by initial distress rank at the outset of the first wave of the program. Vertical lines denote the thresholds where credit size jumps discontinuously.

50 counties. In those settings, clustered standard errors may be biased downward. Therefore, we also calculate p-values using a wild cluster bootstrap (Cameron and Miller, 2015) and report both clustered standard errors and bootstrap p-values.

4.2 Regression discontinuity estimates

The main issue with the DD strategy outlined above is the possibility of correlation between ε_{ct} and the tier 1 variables. If the economic distress ranking were completely random, then counties would be assigned subsidy amounts randomly and we could compare counties across tiers. In practice, the distress rank is weakly correlated with economic variables. If counties that get assigned into tier 1 have systematically worse unobservables that imply different trajectories of employment and unemployment even in absence of the program, then our estimates will be biased.

We tackle this problem by exploiting the discontinuities in tier assignment based on the economic distress rank. We start from a model in differences analogous to the model in (1). We can estimate the effect of the program in year 1996 on outcomes in year k with the following regression

$$Y_{c,k} - Y_{c,1996} = \beta_{0,k} + \theta_k tier1_{c,1996} + \beta X_{c,1996} + \varepsilon_{c,1996} \quad (2)$$

In this model, the county effects from equation (1) have been differenced out and the time effects are absorbed into the constant $\beta_{0,k}$. The coefficient of interest is θ_k, the effect of a tier 1 level subsidy on the outcome k years later. To address the potential correlation between $\varepsilon_{c,1996}$ and $tier1_{c,1996}$, we could estimate a cross-sectional discontinuity regression where we include a function of the distress rank as a regressor. We can repeat this cross sectional estimation strategy for every year between 1996 and 2006.

There are three issues with implementing this discontinuity strategy as proposed above. The first issue is that tier 1 assignment was not entirely based on the economic distress rank. From 2000, poverty and population based rules are added as overrides to the formula for tier assignment. We could classify the counties who change tiers because of these overrides as "defiers", and instrument tier 1 status with tier 1 assignment based on the distress rank. This is not entirely satisfactory, since these counties are not defiers per se. Instead, they are assigned treatment based on a different rule. Wong et al. (2013) show that fuzzy discontinuity design estimates may be severely biased in this setting. They recommend excluding units who are assigned based on additional rules, and estimating equation (2) as a sharp discontinuity design using only counties assigned on the basis of the running variable being considered, in this case, the distress rank.

The second issue is the reduced sample size available to estimate each one of these cross-sectional regressions. For the comparison between tier 1 and tier 2, we only have 70 counties available. This limits our ability to estimate a large number of parameters or implement non-parametric estimators. We reduce the number of coefficients to estimate by making three assumptions: a constant treatment effect assumption and two assumptions on the conditional expectation of outcomes given the distress rank that seem consistent with data before the beginning of the program.

If the effect of the program is constant over time (as assumed in the DD estimates), we can take advantage of the repeated execution of the program. Our constant treatment effect assumption is that the effect of the program only depends on the number of years that have passed since the program takes place. Under this assumption, we can pool the yearly cross sectional regression discontinuity analogs of (2) and estimate a single panel regression

$$Y_{c,t+k} - Y_{c,t} = \beta_0 + \beta_t + f(rank_{c,t}) + \theta_k tier1_{c,t} + \beta X_{c,t} + \varepsilon_{c,t} \tag{3}$$

Here, θ_k is the effect of the subsidy after k years. We allow for time effects β_t to allow for level shifts over time and correlation of observations in the same calendar year, and cluster standard errors by county. We also include lagged controls and dependent variables in $X_{c,t}$

Our additional assumptions concern the functional form of $f(rank_{c,t})$. Figure 3 suggests a linear conditional expectation function of changes in the outcomes given a distress rank. Moreover, the functional form of this relationship does not seem to change at the cutoff threshold. We therefore assume that $f(rank_{c,t})$ is linear and remains constant on either side of the assignment cutoff. We do not allow the function to change across years, allowing only level shifts captured by the time dummies β_t.

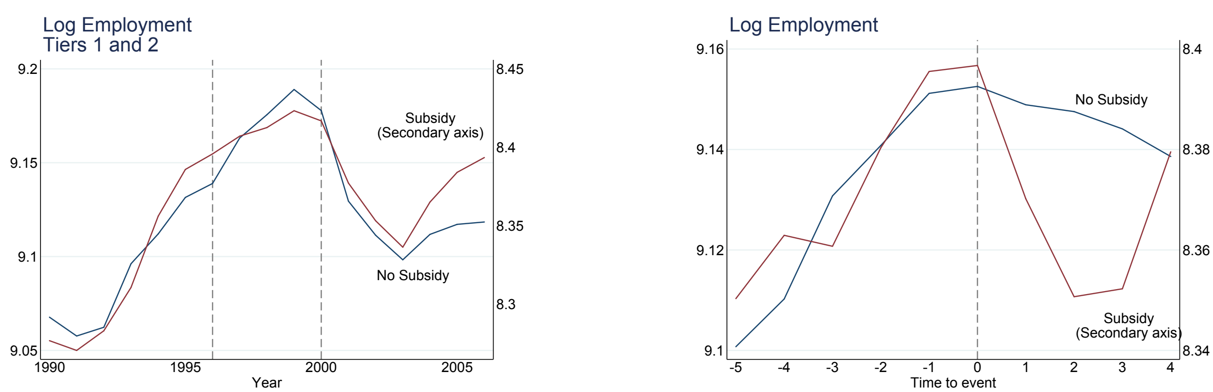
The third issue concerns the dynamic nature of the subsidy program. Consider the effect of the program two years after it is enacted in year t. In year $t+1$, the county would receive the contemporaneous and lagged effect of the program. If the county receives the subsidy in year $t+1$ as well, by year $t+2$ it would experience lagged effects of the program in t and $t+1$ together. Moreover, receiving the program in t may have altered the probability of receiving it in $t+1$. Cellini et al. (2010) show that in this setting, the estimated effects can be interpreted as "Intention to Treat" effects, where employment outcomes are not affected only by the receipt of the subsidy but also by changes in the probability of receiving the subsidy in the future. Our current estimates do not disentangle these channels.

5 Results

Figure 4 portrays the DD analysis for log employment. As can be seen in panel (a), prior to the inception of the program in 1996, subsidy counties - those set to become eligible for the largest credits, had employment growth which was somewhat faster than the control counties, with the control counties catching up afterwards. Since the number of counties eligible for treatment is increasing over time, panel (b) looks at log employment centered on the year of entry to the program. The no subsidy line representing control counties is built by pooling together a different set of control counties for each possible year of treatment entry, weighted by the number of counties entering that year. Although in panel (a) there seems to be a difference in trends between never treated counties and control counties, this trend is not apparent when we compare counties in years leading to program entry. We address some of this trend heterogeneity including lagged dependent variables and county level trends in our specifications. The later RD analysis is designed to be robust to this issue without the need for further assumptions or controls.

Figure 4: Log employment

(a) (b)



Note: Data is from the Bureau of Labor Statistics. Firms in subsidy counties are eligible for the largest size hiring tax credits. Vertical lines denote the initiation of the program in 1996 and the significant expansion in counties eligible for the program in 2000. Left and right axes are adjusted to the same scale.

Figure 5 portrays the DD analysis for unemployment rates. Here pre-trends are more parallel making the difference in differences estimates more credible. A program effect is also apparent here with subsidy county unemployment rates falling after entry into the program, though with a lag.

Table 1 implements the DD strategy for log employment. Column 2 introduces lags

Figure 5: Unemployment rate

(a) (b)

Note: Data is from the Bureau of Labor Statistics. Firms in subsidy counties are eligible for the largest size hiring tax credits. Vertical lines denote the initiation of the program in 1996 and the significant expansion in counties eligible for the program in 2000. Left and right axes are adjusted to the same scale.

of treatment status. Column 3 allows for heterogenous trends at the treatment/control group level. Column 4 allows each county to have its own trend. Column 5 includes lagged time varying county level controls and a lag of the dependent variable. Even once these are added, estimated program impacts are still zero or even negative. Theoretically, the program impact could be zero but not negative. We interpret these estimates as indicating the presence of heterogeneous trends in employment levels for which the differencing strategy cannot account.

Table 2 presents the DD estimates for unemployment. Focusing on column 5, the estimated impact of the program occurs with a lag. Cumulatively, in the fourth year of exposure to the program, unemployment rates are lowered by about 0.6 percentage points.

We now turn to the regression discontinuity estimates. Figure 6 portrays graphical evidence. Counties are arrayed by initial distress rank relative to the threshold where credit size increases. A linear fit in county rank is included, which is constrained to have the same slope on either side of the threshold. Given the prior evidence that program effects appear only with a lag, we focus here on three year differences. Outcomes for counties entering the wave 1 of the program at different points in time are pooled together here. Outcomes are somewhat noisy, but program effects do appear at the thresholds. Panel (a) shows changes in log employment. Counties eligible for the larger subsidies, to the left of the threshold, have about four percentage point higher employment growth. Unemployment rates, shown

Table 1: Difference in difference estimates: Employment

Dependent Variable: Log Employment

	(1)	(2)	(3)	(4)	(5)
Tier1	-0.00398	0.00699	0.00356	-0.00652	0.00181
	(0.0186)	(0.0136)	(0.0118)	(0.00870)	(0.00778)
Lag Tier 1		-0.0225	-0.0239*	-0.0206**	-0.0175*
		(0.0131)	(0.0138)	(0.00904)	(0.00913)
Lag 2 Tier 1		-0.00488	-0.00616	-0.0143	-0.00959
		(0.00750)	(0.00768)	(0.00898)	(0.00774)
Lag 3 Tier 1		0.0208	0.0156	-0.00152	-0.0000123
		(0.0164)	(0.0172)	(0.0139)	(0.0132)
Lag 4 Log population					-0.0242
					(0.453)
Lag 4 Log Real Income per capita					0.187*
					(0.110)
Lag 4 Log Employment					-0.350***
					(0.0720)
Lag 4 Unemployment Rate					-0.00837**
					(0.00353)
Lag 4 Distress Rank					-0.000207
					(0.000262)
Lag 4 Log Labor Force					0.118
					(0.0976)
R^2	0.991	0.994	1.000	1.000	1.000
N	714	588	588	588	546
Counties	42	42	42	42	42
County FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Group trends			Yes		
County trends				Yes	Yes

Clustered standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: Table depicts difference in difference estimates of equation 1. Standard errors are clustered by county. p-values for significance tests are calculated using a wild cluster bootstrap with 500 replications to account for the small number of counties. All columns include county and time effects. Column (4) includes a linear time trend interacted with the Tier 1 dummy. Columns (5) and (6) include linear time trends interacted with county dummies.

Table 2: Difference in difference estimates: Unemployment

Dependent Variable: Unemployment Rate

	(1)	(2)	(3)	(4)	(5)
Tier1	-0.0532	0.264	0.273	0.0724	0.0596***
	(0.283)	(0.187)	(0.187)	(0.164)	(0.148)
Lag Tier 1		-0.0793***	-0.0758	-0.121	-0.0310
		(0.200)	(0.204)	(0.168)	(0.188)
Lag 2 Tier 1		-0.383	-0.380	-0.293	-0.233
		(0.216)	(0.207)	(0.117)	(0.119)
Lag 3 Tier 1		-0.956	-0.943	-0.518	-0.438
		(0.314)	(0.313)	(0.189)	(0.174)
Lag 4 Log population					7.841***
					(4.570)
Lag 4 Log Real Income per capita					0.326***
					(1.825)
Lag 4 Log Employment					2.609
					(1.232)
Lag 4 Unemployment Rate					-0.0988
					(0.0459)
Lag 4 Distress Rank					0.00543***
					(0.00349)
Lag 4 Log Labor Force					-0.211
					(1.245)
R^2	0.652	0.669	0.969	0.982	0.984
N	714	588	588	588	546
$Counties$	42	42	42	42	42
County FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Group trends			Yes		
County trends				Yes	Yes

Clustered standard errors in parentheses

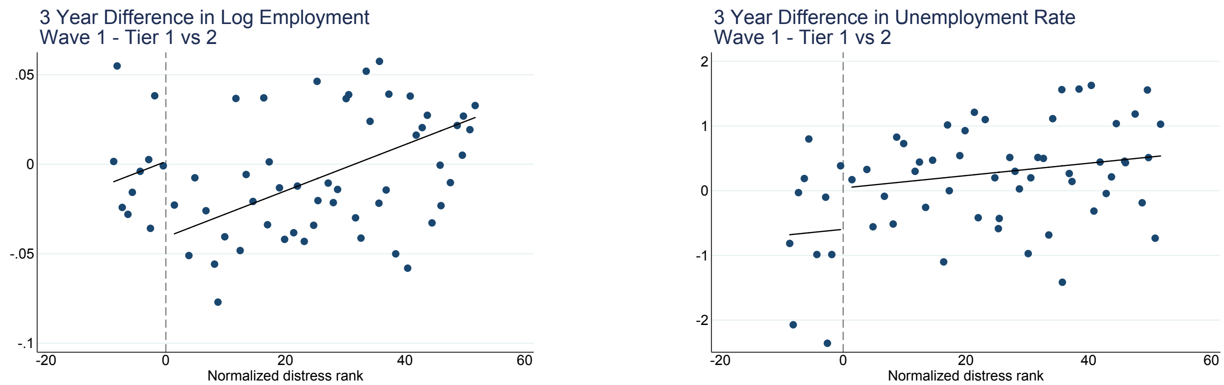
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: Table depicts difference in difference estimates of equation 1. Standard errors are clustered by county. p-values for significance tests are calculated using a wild cluster bootstrap with 500 replications to account for the small number of counties. All columns include county and time effects. Column (4) includes a linear time trend interacted with the Tier 1 dummy. Columns (5) and (6) include linear time trends interacted with county dummies.

in panel (b), are about 0.6 percentage points lower in treated counties. Outcomes are weakly correlated with distress rank, with the assumption of a linear relationship with rank appearing reasonable.

Figure 6: Discontinuities in 3 year differences of employment and unemployment: First wave

(a) (b)



Note: Binned scatterplots of 3-year differences in outcomes. Figure plots the sample mean plus the residuals of a regression of the differenced outcomes on year dummies. Lines depict estimates of the conditional expectation of the outcomes from equation (3). Data is from the Bureau of Labor Statistics. Counties are arrayed by distress rank relative to the threshold. Counties to the left of the threshold are eligible for a larger hiring tax credit.

Table 3 presents the RD estimates for log employment. Pooled differences of one, two, and three years are presented, with even number columns including county level controls for lags of population, income per capita, employment and labor force. At the longer lag shown in column 6, point estimates show an increase in employment for treated counties of 3.1%, though standard errors are large.

Table 4 presents similar RD estimates for unemployment rates. Here program effects show up after a two year lag. After three years, and accounting for controls in column 6, the program is estimated to have reduced unemployment rates by 0.63 percentage points.

The tax credits were limited in the industries which were eligible, with the main sectors being manufacturing, warehousing, wholesale trade, and data processing.[3] Table 5 presents

[3]At the inception of the program in 1996, eligible industries with NAICS codes in parentheses were Manufacturing (31-33), Warehousing (493), Wholesale Trade (42), Research and Development (541710) and Data Processing (Computer Systems Design & Related Services 54151, Software Publishers 511210, Software Reproducing 334611, Data Processing Services 514210, On-Line Information Services 514191). Beginning in 1999, also made eligible were Air Courier Services (492110), Central Administrative Office (551114), Electronic Mail Order (454110), and Customer Service Center (561422).

Table 3: Regression discontinuity estimates: Employment

Dependent Variable: Log Employment

	(1) 1 Year	(2) 1 Year	(3) 2 Years	(4) 2 Years	(5) 3 Years	(6) 3 Years
Tier 1	0.000611	-0.00407	0.0312**	0.0199	0.0463**	0.0309*
	(0.00770)	(0.00749)	(0.0136)	(0.0127)	(0.0186)	(0.0181)
Distress Rank	0.000111	-0.00000460	0.000756**	0.000684*	0.00127**	0.00108**
	(0.000215)	(0.000228)	(0.000365)	(0.000352)	(0.000506)	(0.000532)
R^2	0.099	0.110	0.168	0.205	0.195	0.237
N	406	406	367	367	329	329
Counties	70	70	66	66	66	66
Controls		Yes		Yes		Yes

Clustered standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: Table depicts estimates of equation 3 for 1, 2 and three year differences. Standard errors are clustered by county. All columns include time effects. Columns (2), (4) and (6) include lags of population, real income per capita, log employment and labor force.

Table 4: Regression discontinuity estimates: Unemployment

Dependent Variable: Unemployment Rate

	(1) 1 Year	(2) 1 Year	(3) 2 Years	(4) 2 Years	(5) 3 Years	(6) 3 Years
Tier 1	-0.230	-0.277*	-0.523**	-0.565**	-0.639*	-0.628*
	(0.150)	(0.165)	(0.228)	(0.245)	(0.338)	(0.355)
Distress Rank	0.00429	-0.00103	0.00443	-0.00257	0.00956	0.00249
	(0.00313)	(0.00439)	(0.00493)	(0.00670)	(0.00771)	(0.00966)
R^2	0.485	0.487	0.586	0.587	0.629	0.629
N	406	406	367	367	329	329
Counties	70	70	66	66	66	66
Controls		Yes		Yes		Yes

Clustered standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: Table depicts estimates of equation 3 for 1, 2 and three year differences. Standard errors are clustered by county. All columns include time effects. Columns (2), (4) and (6) include lags of population, real income per capita, log employment and labor force.

RD estimates for aggregate employment within the targeted industries. After three years of program eligibility employment gains in these industries are substantial at about 10%. The same estimations for all non-eligible industries aggregated together are shown in table 6. Point estimates are small negatives though statistically insignificant.

Table 5: Regression discontinuity estimates: Target industry employment

	Dependent Variable: Log Employment - Target Industries					
	(1)	(2)	(3)	(4)	(5)	(6)
	1 Year	1 Year	2 Years	2 Years	3 Years	3 Years
Tier 1	0.00991	0.0105	0.0515	0.0503	0.108**	0.106**
	(0.0211)	(0.0218)	(0.0345)	(0.0354)	(0.0452)	(0.0457)
Distress Rank	0.000364	0.000443	0.000990	0.00101	0.00256**	0.00258**
	(0.000493)	(0.000516)	(0.000821)	(0.000869)	(0.00109)	(0.00117)
R^2	0.044	0.042	0.083	0.084	0.118	0.127
N	384	384	345	345	308	308
Counties	66	66	62	62	62	62
Controls		Yes		Yes		Yes

Clustered standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: Table depicts estimates of equation 3 for 1, 2 and three year differences. Standard errors are clustered by county. All columns include time effects. Columns (2), (4) and (6) include lags of population, real income per capita, log employment and labor force.

As a robustness check, we re-estimate the RD specifications using Census QWI data on hires and separations. The employment gains we attribute to a hiring tax credit program should be stemming from increased hiring and not from decreased separations. Table 7 and table 8 present the RD estimates for hires and separations following a county's eligibility for the program. Though noisy, the estimates are consistent with employment gains resulting from increased hiring rather than decreased separations. In fact, point estimates for separations are actually positive, indicating that some churning of employment may have been induced by the program. If the program did increase the rate of worker turnover at eligible firms in order to generate more tax credits it would not alter the net employment impact of the program, though it would reduce the program's cost effectiveness.

Table 6: Regression discontinuity estimates: Non-target industry employment

Dependent Variable: Log Employment - Non Target Industries

	(1) 1 Year	(2) 1 Year	(3) 2 Years	(4) 2 Years	(5) 3 Years	(6) 3 Years
Tier 1	-0.0103	-0.0128	-0.0000848	-0.00473	-0.0120	-0.0182
	(0.00937)	(0.00914)	(0.0174)	(0.0166)	(0.0279)	(0.0283)
Distress Rank	-0.000151	-0.000194	0.000216	0.000364	0.000321	0.000438
	(0.000209)	(0.000251)	(0.000336)	(0.000382)	(0.000506)	(0.000583)
R^2	0.085	0.080	0.111	0.114	0.139	0.140
N	384	384	345	345	308	308
Counties	66	66	62	62	62	62
Controls		Yes		Yes		Yes

Clustered standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: Table depicts estimates of equation 3 for 1, 2 and three year differences. Standard errors are clustered by county. All columns include time effects. Columns (2), (4) and (6) include lags of population, real income per capita, log employment and labor force.

Table 7: Regression discontinuity estimates: Hires

Dependent Variable: Log Hires - Annual Total

	(1) 1 Year	(2) 1 Year	(3) 2 Years	(4) 2 Years	(5) 3 Years	(6) 3 Years
Tier 1	0.0340*	0.0425*	0.0736**	0.0681*	0.0696	0.0525
	(0.0197)	(0.0241)	(0.0333)	(0.0343)	(0.0512)	(0.0522)
Distress Rank	0.00000396	0.00101*	0.000283	0.00155*	-0.000104	0.00187
	(0.000393)	(0.000537)	(0.000767)	(0.000846)	(0.00115)	(0.00131)
R^2	0.312	0.361	0.387	0.458	0.415	0.508
N	406	406	367	367	329	329
Counties	70	70	66	66	66	66
Controls		Yes		Yes		Yes

Clustered standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: Table depicts estimates of equation 3 for 1, 2 and three year differences. Standard errors are clustered by county. All columns include time effects. Columns (2), (4) and (6) include lags of population, real income per capita, log employment and labor force.

Table 8: Regression discontinuity estimates: Separations

Dependent Variable: Log Separations - Annual Total

	(1) 1 Year	(2) 1 Year	(3) 2 Years	(4) 2 Years	(5) 3 Years	(6) 3 Years
Tier 1	0.00242	0.0131	0.0363	0.0325	0.0317	0.0208
	(0.0140)	(0.0168)	(0.0290)	(0.0288)	(0.0474)	(0.0478)
Distress Rank	-0.0000153	0.000921**	0.0000657	0.00123*	-0.000177	0.00166
	(0.000372)	(0.000425)	(0.000690)	(0.000684)	(0.00110)	(0.00105)
R^2	0.364	0.400	0.485	0.551	0.443	0.523
N	406	406	367	367	329	329
Counties	70	70	66	66	66	66
Controls		Yes		Yes		Yes

Clustered standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: Table depicts estimates of equation 3 for 1, 2 and three year differences. Standard errors are clustered by county. All columns include time effects. Columns (2), (4) and (6) include lags of population, real income per capita, log employment and labor force.

6 Discussion

Hiring tax credits are a popular tool implemented at various times in many U.S. states as a way to lure businesses or revitalize moribund local economies. Assessing their efficacy is challenging though as their implementation is typically expressly endogenous to local conditions and expected future prospects. We make use of the unusual institutional features of a program in the state of North Carolina to get causal estimates of the impact of hiring tax credits on employment and unemployment rates. We document the importance of accounting for unobservables and mean reversion which will bias difference in differences estimates. Our RD estimates show a boost to employment from the program of around 3%. We also find substantial impacts on unemployment rates, with treated counties experiencing about 0.6 percentage point lower unemployment rates than a counterfactual. Scaling by the difference in subsidy size across the policy threshold, this corresponds to a 0.7 percentage point reduction in unemployment rates for a $10,000 credit.[4]

[4]Tier 1 credit size is $12,500, or $9,000 more than the next tier which averages $3,500.

References

Amior, Michael and Alan Manning (2015), "The Persistence of Local Joblessness." CEP Discussion Papers dp1357, Centre for Economic Performance, LSE.

Bartik, Timothy J. (2001), *Jobs for the Poor: Can Labor Demand Policies Help?* W.E. Upjohn Institute for Employment Research.

Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan (2004), "How much should we trust differences-in-differences estimates?" *The Quarterly Journal of Economics*, 119, 249–275.

Busso, Matias, Jesse Gregory, and Patrick Kline (2013), "Assessing the Incidence and Efficiency of a Prominent Place Based Policy." *American Economic Review*, 103, 897–947.

Cameron, Collin A. and Douglas L. Miller (2015), "A practitioner's guide to cluster-robust inference." *Journal of Human Resources*, 50, 317–372.

Cellini, Stephanie Riegg, Fernando Ferreira, and Jesse Rothstein (2010), "The Value of School Facility Investments: Evidence from a Dynamic Regression Discontinuity Design." *The Quarterly Journal of Economics*, 125, 215–261.

Chirinko, Robert S. and Daniel J. Wilson (2016), "Job Creation Tax Credits, Fiscal Foresight, and Job Growth: Evidence from U.S. States." CESifo Working Paper Series 5771, CESifo Group Munich.

Department of Commerce (2014), "Measuring economic distress in north carolina." Technical report.

Freedman, Matthew (2013), "Targeted Business Incentives and Local Labor Markets." *Journal of Human Resources*, 48, 311–344.

Heckman, James J., Robert J. Lalonde, and Jeffrey A. Smith (1999), "The economics and econometrics of active labor market programs." In *Handbook of Labor Economics* (O. Ashenfelter and D. Card, eds.), volume 3, chapter 31, 1865–2097, Elsevier.

Kline, Patrick and Enrico Moretti (2013), "Place Based Policies with Unemployment." *American Economic Review*, 103, 238–43.

Neumark, David (2013), "Spurring job creation in response to severe recessions: Reconsidering hiring credits." *Journal of Policy Analysis and Management*, 32, 142–171.

Neumark, David and Diego Grijalva (2015), "The Employment Effects of State Hiring Credits." IZA Discussion Papers 9146, Institute for the Study of Labor (IZA).

Program Evaluation Division (2015), "Final report to the joint legislative program evaluation oversight committee." Technical Report 2015-11, North Carolina General Assembly.

Wong, Vivian C, Peter M Steiner, and Thomas D Cook (2013), "Analyzing regression-discontinuity designs with multiple assignment variables: A comparative study of four estimation methods." *Journal of Educational and Behavioral Statistics*, 38, 107–141.